

Enhanced Memory Retrieval Capacity in Hopfield Neural Networks

Do-Hyun Kim,^{1,*} Jinha Park,² and B. Kahng^{2,†}

¹*Department of Physics, Sogang University, Seoul 04107, Korea*

²*CCSS, CTP and Department of Physics and Astronomy, Seoul National University, Seoul 08826, Korea*

(Dated: August 4, 2016)

The Hopfield model is a paradigmatic model for understanding neural activities associated with memory retrieval in view of the collective behavior of neurons. The analytical solution of the model on fully connected networks revealed that information stored in the memory can be retrieved without any error up to a certain threshold of storage capacity, beyond which the information is completely lost. A normal human brain, however, has imperfect memory with some error even when its storage capacity is small. Moreover, it has not yet been proven whether such a storage capacity threshold exists. Here, we show analytically that as the network changes from a hub-damaged network (e.g., shown in brains of schizophrenia or comatose patients) to a scale-free network (e.g., shown in brains of ordinary people) with degree exponent two (close to the degree exponent value of the networks constructed from fMRI experimental data), the memory retrieval capacity is tremendously enhanced, but with some error. These features seem to be in accordance with what we experience from our ordinary brain and to a recent discovery of enormously high storage capacity in human brain. Our novel model may be useful for designing realistic artificial neural networks.

PACS numbers: 64.60.aq, 75.10.Nr, 84.35.+i, 89.75.Da

I. INTRODUCTION

Human neuroscience has garnered increasing attention in recent years as functional magnetic resonance imaging (fMRI) experiments have become more easily accessible and their results have found immediate application in cognitive studies. Data obtained from fMRI or anatomical experiments allow us to construct coarse-grained neural networks associated with given neural activities. This advanced technique provides new insight into understanding of the mechanisms and efficiency of the diverse neural activities of our brain network. One such subject is the retrieval of associative memory of neural networks [1]. New memories are initially labile and can be stabilized into long-term memories through retrieval and re-consolidation processes. In neurobiology, the memory retrieval process has been investigated from the perspective of local protein synthesis. However, it has been proposed in statistical physics that memory retrieval in our brain occurs by way of pattern recognition of collective excitations of associated neurons. In this regard, a spin model, the Hopfield model [2] introduced a long time ago has since become a paradigmatic model for memory retrieval in neural networks.

The Hopfield model was analytically solved for fully connected networks, known as the mean-field solution [3–5]. Recently, however, fMRI and anatomical experiments have revealed that human neural networks are scale-free (SF) networks [6, 7]. This means that the number of synapses of each neuron, referred to as the degree in graph theory, is heterogeneous and their distribution follows the power law $P_d(k) \sim k^{-\gamma}$, where k denotes the

degree and γ is the degree exponent, or a heavy-tailed distribution [8]. Particularly, we notice that the degree exponent obtained from the fMRI data is close to $\gamma = 2$ [9–11].

It is well known that the functionality of an individual neuron is partly determined by the pattern of its anatomical connections with other neurons in the brain network, so that neurons with similar connection patterns tend to exhibit similar functionality [6]. It is thus possible to infer that such fMRI results on the functional connectivities reflect the structural SF feature of the brain networks on the neuron-to-neuron scale. For instance, the brains of schizophrenia [12, 13] or comatose patients [14] may be related to the malfunction of the hub of heterogeneous networks on the neuron-to-neuron scale. Therefore, it is expected that a Hopfield neural network model structured on SF networks composed of numerous neuron vertices may be very helpful to understand some specific features of realistic brain networks.

This paper aims to investigate the properties of retrieval patterns created by the Hopfield model on SF networks in the limit $\gamma \rightarrow 2$ using analytic method and to compare the obtained results with previous numerical results obtained from the fully connected networks and diverse SF networks such as the Barabási-Albert model with $\gamma = 3$ [15] and the Molloy-Reed model with several values of γ but in extremely narrow range of storage capacity a [16]. We also use the Chung-Lu model [17, 18] for constructing uncorrelated SF networks over the entire range of γ . The details of the Chung-Lu model are presented in Appendix A.

Using the replica method, we analytically obtain the phase diagram of the Hopfield model for various degree exponent values (Fig. 1): When $\gamma < \gamma_c \simeq 2.04$, the retrieval phase spans most of the low-temperature region, thus the system is enhanced for memory retrieval. At

* dohyunkim@sogang.ac.kr

† bkahng@snu.ac.kr

$T = 0$, error rate is obtained as a function of storage capacity a (Fig. 2). Remarkably for SF networks with $\gamma = 2.01$, our result shows that the error rate is almost zero when the number of stored memories p is a few, and gradually increases but is less than 0.3 even when p is increased up the total number of nodes in a network. This implies that the storage capacity is unlimited in the thermodynamic limit, but with some error. In short, our findings suggest that the memory retrieval capacity is tremendously enhanced in the SF network, but with some error. These features seem to be in accordance with what we experience from our ordinary brain and to a recent discovery of enormously high storage capacity in human brain [20].

II. OUTLINE OF ANALYTIC SOLUTIONS

In the Hopfield model, each neuron in a neural network is regarded as an Ising spin with spin states representing excited and rest states for transmitting or not transmitting a signal, respectively [2]. Memory is regarded as a collective pattern of neuron excitations. An excitation pattern of a neural network is denoted by $\{\xi_i^\mu\}$, where i is the index of the neuron, $\mu (= 1, \dots, p)$ is the index of the excitation pattern, and p is the number of existing patterns, equivalent to the number of stored memories. The value of ξ_i^μ is either +1 or -1. The state of the Ising spin at each neuron is updated with time as

$$S_i(t+1) = \text{sgn}\left(\sum_j J_{ij} S_j(t)\right), \quad (1)$$

such that $S_i(t+1)$ is affected by the states of functionally connected neurons at time t . The coupling strength J_{ij} , known as the synapse efficacy, for SF network is defined as

$$J_{ij} = \frac{1}{K} \sum_{\mu=1}^p \xi_i^\mu \xi_j^\mu \quad (2)$$

following Hebb's rule, where K is the mean degree of a given Chung-Lu SF network G of size N . The time evolution is explored using the Hamiltonian

$$\mathcal{H} = - \sum_{(i,j) \in G} J_{ij} S_i S_j, \quad (3)$$

where the i -th spin value, i.e., the state of i -th neuron, is assumed as $S_i = \pm 1$, and J_{ij} is assumed symmetric $J_{ij} = J_{ji}$, and J_{ij} is set zero for any disconnections in G . The probability for a given graph G to exist over ensemble is given as

$$P_K(G) = \prod_{(i,j) \in G} f_{ij} \prod_{(i,j) \notin G} (1 - f_{ij}) \quad (4)$$

where $f_{ij} = 1 - \exp(-NKw_i w_j)$. Then the graph ensemble average for a given physical quantity A is taken

as

$$\langle A \rangle_K = \sum_G P_K(G) A(G), \quad (5)$$

where $\langle \dots \rangle_K$ denotes the average over different graph configurations.

The independent random distributed variable of μ -th pattern of memory, $\xi_i^\mu = \pm 1$, follows the probability distribution

$$P(\{\xi_i\}) = \prod_{i \in G} \left[\frac{1}{2} \delta(\xi_i^\mu - 1) + \frac{1}{2} \delta(\xi_i^\mu + 1) \right]. \quad (6)$$

Then the ensemble average for a given physical quantity, A is taken as

$$\langle A \rangle_\xi = \int d\xi_i P(\{\xi_i\}) A(\{\xi_i\}), \quad (7)$$

where $\langle \dots \rangle_\xi$ is an average over the quenched disorder of ξ_i . A retrieval pattern is recognized as a spin state at the free energy minimum. In this approach, temperature T is involved, which represents the noise strength from a stochastic process of the system.

The replica method is conventionally used to evaluate the free energy [21, 22]. Using the replica trick, the free energy F is evaluated as $-\beta F = \langle \ln Z \rangle_\xi / K = \lim_{n \rightarrow 0} [\langle \langle Z^n \rangle_\xi \rangle_K - 1] / n$, where Z is the partition function for a given distribution of $\{\xi_i\}$ on a particular graph G and $\beta = 1/T$. By evaluating the n -th power of the partition function $\langle \langle Z^n \rangle_\xi \rangle_K$, we obtain the free energy as shown in Eq. (B18) in Appendix B. For simplicity, only the first pattern ($\mu = 1$) is retrieved. This simplification is based on the fact that the replica symmetry only appears to be broken in a particular case at $T = 0$ [5]. The mathematical details can be found in Appendix C.

We characterize the phases of the Hopfield model by three order parameters defined as follows: i) The overlap parameter defined as $m_\alpha^\mu \equiv \sum_i w_i \langle \xi_i^\mu S_i^\alpha \rangle$, which represents the extent to which the μ -th pattern of memory ξ^μ and the α -th state of the system S^α overlap with each other. This quantity is used to measure the retrieval success rate of the μ -th memory. ii) The spin glass order parameter $q_{\alpha\beta} \equiv \sum_i w_i \langle S_i^\alpha S_i^\beta \rangle$ representing the extent to which the two states α and β of the replica overlap each other. iii) A new quantity is introduced as $r_{\alpha\beta} \equiv (N/p) \sum_{\mu=2}^p m_\alpha^\mu m_\beta^\mu$. Henceforth, we denote p/N as the storage capacity a , because we are interested in the case $p \sim \mathcal{O}(N)$. This quantity is necessary to derive the first two order parameters, and is interpreted as the sum of the effects of each non-retrieval pattern. We remark that the overlap and the spin-glass order parameters are defined similarly to those of the mean-field solution, but the factor w_i appears, which is required to take into account heterogeneous degrees. w_i is proportional to the degree k_i of neuron i [23–26]. Next, we take the replica-symmetric solution by setting $q_{\alpha\beta} = q$ and $r_{\alpha\beta} = r$ for all $\alpha \neq \beta$ and $m_\alpha^1 = m$ for all α . Explicit formulas for those parameters are presented in Appendix C.

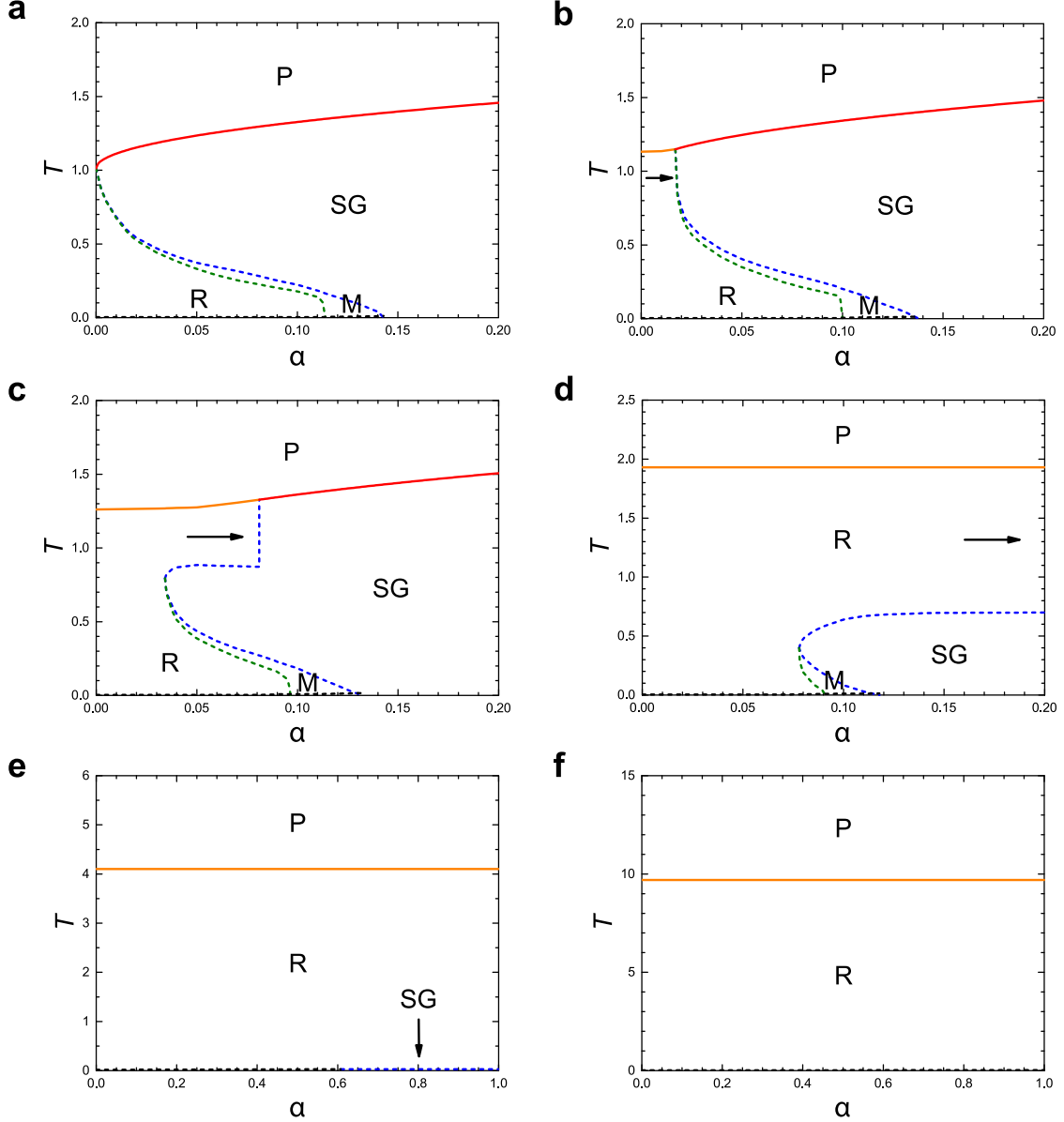


FIG. 1. **Phase diagram of the Hopfield model in the plane of (T, a) .** Here T denotes temperature and a the storage capacity. Degree exponent γ is **a**, ∞ , **b**, 5.0, **c**, 4.0, **d**, 3.0, **e**, 2.04, and **f**, 2.01. **P** denotes a paramagnetic phase, **SG** a spin glass phase, **M** a mixed phase, and **R** a retrieval phase. As the degree exponent γ is decreased from **a**, ∞ (the ER limit) through **f**, $\gamma = 2.01$, the R phase not only intrudes into the region of the SG phase, but also raises the boundary of the P phase to a higher temperature region. Eventually the SG phase remains on the axis $T = 0$ when $\gamma = \gamma_c \simeq 2.04$, in which case the R phase spans most of the low-temperature region. Thus, the system is enhanced for memory retrieval. The phase boundary was obtained by performing numerical calculations for the Chung-Lu SF networks with the system size $N = 1000$ and mean degree $K = 5.0$. Solid and dotted curves indicate second-order and first-order transitions, respectively. Note that the dotted black line of each panel near zero temperature represents the Almeida-Thouless line [19]. Thus, the replica-symmetric solution is valid over almost the entire region.

III. PHASE DIAGRAM AND ERROR RATES

The phase diagram we obtained are as follows: Fig. 1 shows the phase diagrams of the Hopfield model in the $T - a$ plane for different values of γ of SF networks with

$N = 1000$ and $K = 5.0$. First, we consider the case of the Erdős and Rényi (ER) network, equivalent to the limit $\gamma \rightarrow \infty$. The phase diagram shown in Fig. 1(a) is similar to that in Ref. [4], obtained from a fully connected network. Four different phases exist: i) The paramagnetic

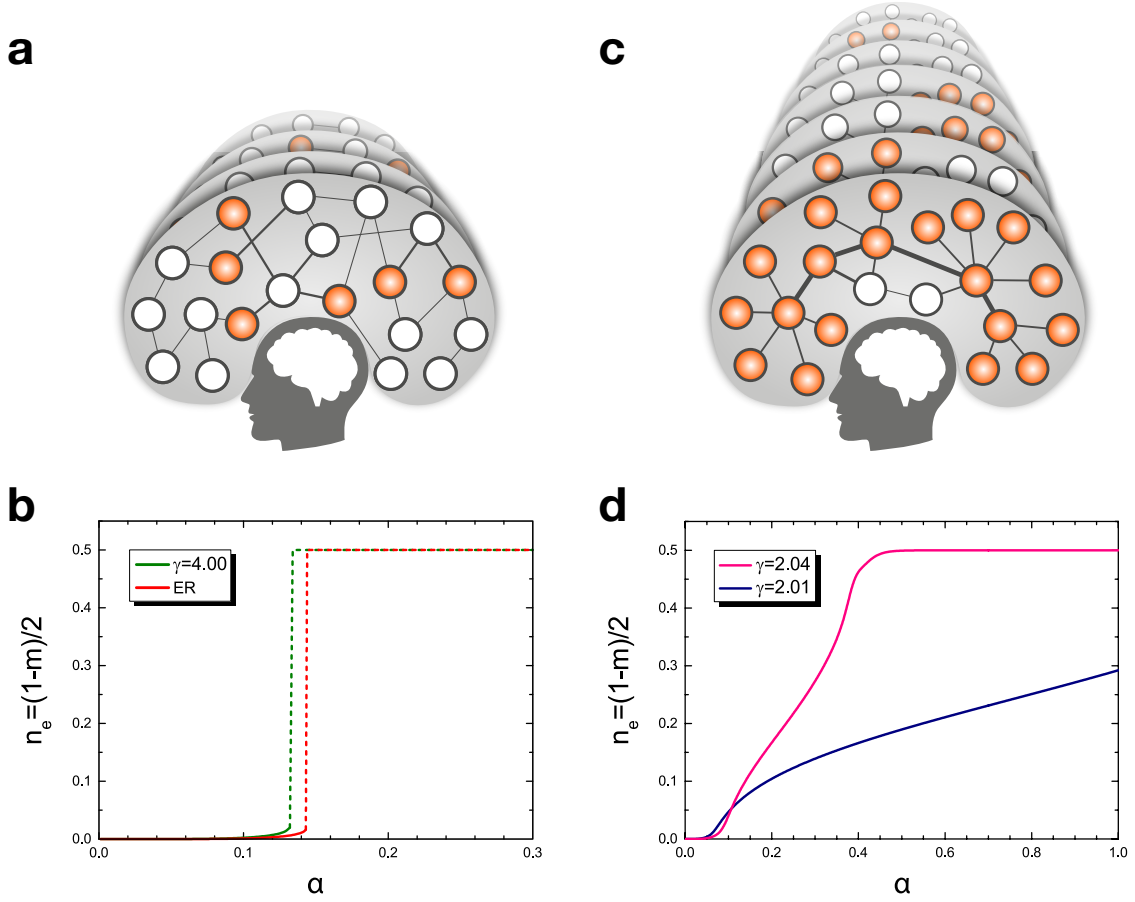


FIG. 2. **Conceptual figures of the storage capacities and the error rates.** **a**, for an ER random network and **c**, for a SF network. **b**, and **d**, Plot of the error rate $n_e \equiv (1 - m)/2$ vs storage capacity a for several γ values of the Chung-Lu model at $T = 0$. Here, numerical values are obtained using $N = 1000$ and $K = 5.0$. The dotted lines for $\gamma \gg 2.0$ indicate the sudden jumps from small error rates to the state of $n_e = 0.5$.

phase denoted as P, in which $m = 0$, $q = 0$, and $r = 0$ because of thermal fluctuations. ii) The spin-glass phase denoted as SG, in which $m = 0$, $q > 0$, and $r > 0$. In this phase, the retrieval of stored patterns is not possible. Thus, it is often referred to as the confusion phase. iii) The retrieval phase denoted as R, in which $m > 0$, $q > 0$, and $r > 0$, and in which the retrieval of stored memory is possible. Finally, the mixed phase denoted as M, in which the patterns of both the retrieval and the spin-glass phases coexist.

The free energy of the system in each phase has also been investigated in Ref. [4]. In phase R, a stored pattern, for instance, $\{\xi_i^\mu\}$ is matched to a state of the system $\{S_i^\alpha\}$ at the global minimum of free energy. However, in phase M, a stored pattern is matched to a state of the system in a metastable state, while the SG state lies at the global minimum.

For $a = 0$, which occurs when $p \sim \mathcal{O}(N)$ in the limit $N \rightarrow \infty$, there exist two phases, P and R. As a is increased slightly from $a = 0$, the SG and M phases appear between the phases P and R. The transition between the

P and the SG phases is a second-order transition, whereas that between SG and M is a first-order transition. Likewise, the transition between M and R is also first order. At $T = 0$, the transition between R and M occurs at $a_m \simeq 0.114$, and the transition between M and SG occurs at $a_c \simeq 0.143$. Therefore, as temperature is lowered from a sufficiently high value, successive transitions occur following the steps $P \rightarrow SG \rightarrow M \rightarrow R$ for $a < a_m$. For $a_m < a < a_c$, successive transitions occur following the steps $P \rightarrow SG \rightarrow M$, and for $a > a_c$, a transition from $P \rightarrow SG$ occurs.

Second, we obtain the phase diagram for finite γ values. The phase diagram in Fig. 1(b)-(e) undergoes drastic changes depending on γ . Remarkably, the R phase intrudes into the region of the SG phase, but it also raises the boundary of the P phase to a higher-temperature region. As γ approaches 2.04 in Fig. 1(e), we observe that the R phase prevails, whereas the SG phase shrinks until it only exists at $T = 0$ in the region $a > 0.6$. Such changes in the $T - a$ diagrams can be understood analytically by examining the phase boundary between the P

and SG phases. The details are presented in Appendix D.

Third, we consider a particular case, the noiseless case $T = 0$. The order parameters m , q and r are derived in Appendix C. Fig. 2 shows the behavior of the error rate $n_e \equiv (1 - m)/2$ as a function of a on SF networks with $N = 1000$ and $K = 5.0$. This figure is obtained analytically for various degree exponent values including $\gamma \simeq 2.0$. We first consider the case in which $\gamma \rightarrow \infty$, i.e., the ER limit. The result of this case is almost the same as that obtained in Ref. [4]. When a is less than $a_c \simeq 0.143$, n_e is very small, such that the error rate is negligible and the system is almost in the error-free state. The obtained value a_c is close to $a_c \simeq 0.138$, which is obtained from the fully connected network in Ref. [4]. As a is increased and reaches a_c , n_e suddenly jumps to 0.5. This means that for $a > a_c$, the system is in the error-full state (i.e., the state of complete confusion). As γ decreases, this behavior persists up to $\gamma_c \simeq 2.7$, and no longer holds for $\gamma \leq \gamma_c$. Note that the value of a_c slightly decreases with decreasing γ , but their dependences are almost negligible.

Next, when γ approaches 2.0, n_e noticeably changes in its form from the step-function-like shape to a monotonously increasing one. As γ is lowered further and approaches 2.0, the range of a for the state of complete confusion, with $n_e = 0.5$, disappears and a_c cannot be defined anymore. For instance, at $\gamma = 2.01$, the error rate becomes less than 0.3 for the entire range of a . Therefore, when γ is lowered to as small as 2.0, while the range of a for the state of $n_e = 0$ (the state of perfect retrieval without error) is reduced, the range of a for the state of $n_e = 0.5$ (the state of complete confusion) disappears. The details are presented in Appendix E. These behaviors have never been observed yet in previous studies [27].

IV. DISCUSSION AND SUMMARY

Our results might provide some insight into constructing a certain artificial neural network (ANN): If we want to construct an ANN that contains a function of perfect memory recall with a large value of a_c as a basic topology of the ANN, it may be more appropriate to choose an ER-type random network or a fully connected one. However, should we prefer to construct an ANN that sustains a large range of storage capacity, but only allows a small range of errors, it may be more appropriate to choose an SF-type network with $\gamma \simeq 2.0$.

Our main results imply that as the network changes from a hub-damaged network to a scale-free network with degree exponent two, the memory retrieval capacity is tremendously enhanced, but with some error. These features seem to be in accordance with what we experience from our ordinary brain and to a recent discovery of enormously high storage capacity in human brain. Thus, hubs, i.e., neurons with a large number of synapses, and other nodes with heterogeneous degrees in neural networks, play a central role in enhancing the

capacity of memory retrieval. It is interesting to notice that a normal human brain has such a structure, even though detailed structural properties such as modularity and degree correlation are different. We consider the effect of degree correlation near $\gamma = 2.0$ by performing similar analysis for the static model [23–26] of which the degree correlation is disassortative between $2 < \gamma < 3$. We find that the correlated SF networks in the vicinity of $\gamma = 2.0^+$ have lower values of n_e compared with those of the uncorrelated ones. The details are presented in Appendix F. Finally, we have checked the error rate n_e on several real neural networks, of which the degree distributions are heavy-tailed, but the degree exponent cannot be determined precisely due to the small system sizes of those networks. The obtained patterns of the error rates are indeed similar to what we predicted theoretically. The details are presented in Appendix G.

ACKNOWLEDGEMENTS:

This work was supported by the NRF of Korea (Grant Nos. 2014R1A3A2069005 and 2015R1A5A7037676) and Sogang University (Grant No. 201610033.01). We would like to express our appreciation toward S.-H. Lee and Joonwon Lee.

Appendix A: Construction of scale-free networks

To construct uncorrelated SF networks, we use the Chung-Lu (CL) model [17, 18]: We start with a fixed number of N vertices. Each vertex i ($i = 1, 2, \dots, N$) is assigned a weight

$$w_i = \frac{(i + i_0 - 1)^{-\mu}}{\sum_{j=1}^N (j + i_0 - 1)^{-\mu}}, \quad (\text{A1})$$

where μ is a control parameter in the range $[0, 1)$, and i_0 is constant given by

$$i_0 = \begin{cases} [10\sqrt{2}(1 - \mu)]^{1/\mu} \cdot N^{1-1/2\mu} & (1/2 < \mu < 1), \\ 1 & (0 \leq \mu < 1/2). \end{cases} \quad (\text{A2})$$

A pair of vertices (i, j) is chosen with the probabilities w_i and w_j , respectively, and they are connected with an edge, unless the pair is already connected. This process is repeated $NK/2$ times. Then, the resulting network becomes an uncorrelated SF network following a power-law degree distribution, $P_d(k) \sim k^{-\gamma}$, where k denotes the degree and γ is the degree exponent with $\gamma = 1 + 1/\mu$. In such random networks, the probability that a given pair of vertices (i, j) ($i \neq j$) is not connected by an edge, as denoted by $1 - f_{ij}$, is given by $(1 - 2w_i w_j)^{NK/2} \simeq \exp(-NKw_i w_j)$, while the connection probability $f_{ij} = 1 - \exp(-NKw_i w_j)$.

As a particular case, when we choose i_0 as 1 for all the values of μ , the CL model reduces to the static model [23–26], which has correlations for the range of $1/2 < \mu < 1$

[23, 24]. Therefore, the weights for the static model are given by

$$w_i = \frac{i^{-\mu}}{\zeta_N(\mu)} \quad (\text{A3})$$

where μ is a control parameter in the range $[0, 1)$, and $\zeta_N(\mu) \equiv \sum_{j=1}^N j^{-\mu} \simeq N^{1-\mu}/(1-\mu)$. Note that $f_{ij} \simeq NKw_iw_j$ for finite K , however, $f_{ij} \simeq 1$ for $2 < \gamma < 3$ and $ij \ll N^{3-\gamma}$.

For the Erdős-Rényi (ER) graph [28–30], μ becomes 0 and the weights of both models become $w_i = 1/N$, independent of the index i . Since $w_iw_j = 1/N^2$, the fraction of bonds present becomes $f_{ij} \simeq K/N$ and the total number of the connected edges L is $NK/2$. So K becomes the mean degree in the ER graph.

When K approaches N , this network becomes a fully-connected one or a regular lattice with infinite-range interaction. Note that previous studies on the Hopfield model focused mainly on such extreme cases of $K \rightarrow N$ [3–5, 21, 22].

Appendix B: Solution of the Hopfield model on scale-free networks

Now we consider the Hopfield model on SF networks. Since Amit *et al.* successfully applied replica analysis of spin glass theory to the Hopfield model [3–5], neural networks have been regarded as analogous systems of spin glasses (SGs) [21, 22]. The SG transitions in Euclidean space have already been studied by means of various theoretical methods [21, 22]. Most of such studies have been carried out on regular lattices or for the infinite-range interaction model on fully connected graphs. To study the SG transitions on SF networks, we followed the previous approach used for the dilute Ising SG model with infinite-range interactions, i.e., the Ising SG model on the ER graph, first performed by Viana and Bray [31–36], and applied it successfully to the static model of SF networks [25]. The Ghatak-Sherrington SG model was also studied on the same networks [26].

Here, the replica method is used to evaluate the free energy, i.e., $-\beta F = \lim_{n \rightarrow 0} [\langle \langle Z^n \rangle_\xi \rangle_K - 1]/n$. To proceed, we evaluate the n -th power of the partition function Z^n ,

$$\begin{aligned} \langle \langle Z^n \rangle_\xi \rangle_K &= \text{Tr} \left\langle \left\langle \exp \left(\frac{\beta}{K} \sum_{(ij) \in G} \sum_{\mu=1}^p \sum_{\alpha}^n \xi_i^\mu \xi_j^\mu S_i^\alpha S_j^\alpha \right) \right\rangle_\xi \right\rangle_K \\ &= \text{Tr} \exp \left[\sum_{i < j} \ln \left\{ 1 + f_{ij} \left(\left\langle \exp \left(\frac{\beta}{K} \sum_{\mu=1}^p \sum_{\alpha}^n \xi_i^\mu \xi_j^\mu S_i^\alpha S_j^\alpha \right) \right\rangle_\xi - 1 \right) \right\} \right], \end{aligned} \quad (\text{B1})$$

where the trace Tr is taken over all replicated spins $S_i^\alpha (=$

$\pm 1)$, and $\alpha = 1, \dots, n$ is the replica index. Using the relation,

$$\left\langle \exp \left(\frac{\beta}{K} \sum_{\mu=1}^p \sum_{\alpha}^n \xi_i^\mu \xi_j^\mu S_i^\alpha S_j^\alpha \right) \right\rangle_\xi = \left\langle \prod_{\mu=1}^p \prod_{\alpha}^n \left[\cosh(\beta/K) \left(1 + \xi_i^\mu \xi_j^\mu S_i^\alpha S_j^\alpha \tanh(\beta/K) \right) \right] \right\rangle_\xi, \quad (\text{B2})$$

$\langle \langle Z^n \rangle_\xi \rangle_K$ term becomes

$$\begin{aligned} \langle \langle Z^n \rangle_\xi \rangle_K &\stackrel{n \rightarrow 0}{=} \text{Tr} \prod_{i \neq j} \exp \left[\frac{1}{2} \left\langle \left[NKw_iw_j + NK\mathbf{T}_1 w_i \xi_i^1 S_i^\alpha w_j \xi_j^1 S_j^\alpha + \dots \right] \prod_{\mu=2}^p \left[1 + \mathbf{T}_1 \xi_i^\mu S_i^\alpha \xi_j^\mu S_j^\alpha + \dots \right] \right\rangle_\xi - NKw_iw_j \right] \\ &= \text{Tr} \prod_{i \neq j} \exp \frac{1}{2} \left\langle NK\mathbf{T}_1 (w_i \xi_i^1 S_i^\alpha) (w_j \xi_j^1 S_j^\alpha) + \sum_{\mu=2}^p NK\mathbf{T}_1 (w_i \xi_i^\mu S_i^\alpha) (w_j \xi_j^\mu S_j^\alpha) \right. \\ &\quad \left. + NK\mathbf{T}_2 (w_i \xi_i^1 S_i^\alpha \xi_i^1 S_i^\beta) (w_j \xi_j^1 S_j^\alpha \xi_j^1 S_j^\beta) + \sum_{\mu=2}^p NK\mathbf{T}_2 (w_i \xi_i^\mu S_i^\alpha \xi_i^\mu S_i^\beta) (w_j \xi_j^\mu S_j^\alpha \xi_j^\mu S_j^\beta) + \dots \right\rangle_\xi, \end{aligned} \quad (\text{B3})$$

where

$$\begin{aligned} \mathbf{T}_l(T) &\equiv \cosh^n(\beta/K) \tanh^l(\beta/K) \xrightarrow{n \rightarrow 0} \tanh^l(\beta/K) \\ (l &= 1, 2, \dots). \end{aligned} \quad (\text{B4})$$

It will be assumed for simplicity that only the first pat-

tern ($\mu = 1$) to be retrieved. Therefore, it is convenient to separate the contribution of the first pattern from the

rest. The $\langle\langle Z^n \rangle_\xi\rangle_K$ term can be linearized using the Hubbard-Stratonovich transformation

$$\exp\left\{\frac{1}{2}\lambda a^2\right\} = \sqrt{\frac{\lambda}{2\pi}} \int_{-\infty}^{\infty} dx \exp\left\{-\frac{1}{2}\lambda x^2 + \lambda a x\right\}. \quad (\text{B5})$$

Therefore,

$$\begin{aligned} \langle\langle Z^n \rangle_\xi\rangle_K &= \text{Tr} \int \prod_{\mu}^p \prod_{\alpha}^n dm_{\alpha}^{\mu} \left\langle \exp \left[NK\mathbf{T}_1 \sum_{\alpha}^n \left\{ -\frac{1}{2}(m_{\alpha}^1)^2 + \sum_i^N m_{\alpha}^1(w_i \xi_i^1 S_i^{\alpha}) \right\} \right] \right. \\ &\quad \times \exp \left[NK\mathbf{T}_1 \sum_{\mu=2}^p \sum_{\alpha}^n \left\{ -\frac{1}{2}(m_{\alpha}^{\mu})^2 + \sum_i^N m_{\alpha}^{\mu}(w_i \xi_i^{\mu} S_i^{\alpha}) \right\} \right] \cdots \left. \right\rangle_{\xi} \\ &= \text{Tr} \int \prod_{\mu}^p \prod_{\alpha}^n dm_{\alpha}^{\mu} \exp \left[NK\mathbf{T}_1 \sum_{\alpha}^n \left\{ -\frac{1}{2}(m_{\alpha}^1)^2 + \sum_i^N m_{\alpha}^1(w_i \xi_i^1 S_i^{\alpha}) \right\} \right] \\ &\quad \times \exp \left[NK\mathbf{T}_1 \sum_{\mu=2}^p \sum_{\alpha}^n \left\{ -\frac{1}{2}(m_{\alpha}^{\mu})^2 \right\} + \sum_{\mu=2}^p \sum_i^N \ln \cosh \left(NK\mathbf{T}_1 \sum_{\alpha}^n m_{\alpha}^{\mu}(w_i S_i^{\alpha}) \right) \right] \cdots. \end{aligned} \quad (\text{B6})$$

In the thermodynamic limit ($N \rightarrow \infty$), we expand

$\ln \cosh(\cdots)$ to the second order of m_{α}^{μ} , so that

$$\begin{aligned} &\exp \left[NK\mathbf{T}_1 \sum_{\mu=2}^p \sum_{\alpha}^n \left\{ -\frac{1}{2}(m_{\alpha}^{\mu})^2 \right\} + \sum_{\mu=2}^p \sum_i^N \ln \cosh \left(NK\mathbf{T}_1 \sum_{\alpha}^n m_{\alpha}^{\mu}(w_i S_i^{\alpha}) \right) \right] \\ &\simeq \exp \left[-\frac{1}{2} NK\mathbf{T}_1 \sum_{\mu=2}^p \sum_{\alpha\beta}^n m_{\alpha}^{\mu} K_{\alpha\beta} m_{\beta}^{\mu} \right], \end{aligned} \quad (\text{B7})$$

where $K_{\alpha\beta} \equiv \delta_{\alpha\beta} - K\mathbf{T}_1 \sum_i^N w_i S_i^{\alpha} S_i^{\beta}$. Since Eq.(16) is

quadratic in m_{α}^{μ} , the integral over m_{α}^{μ} can be carried out using of the multi-variable Gaussian integral as follows:

$$\begin{aligned} &\int \prod_{\mu=2}^p dm_{\alpha}^{\mu} \exp \left[-\frac{1}{2} NK\mathbf{T}_1 \sum_{\mu=2}^p \sum_{\alpha\beta}^n m_{\alpha}^{\mu} K_{\alpha\beta} m_{\beta}^{\mu} \right] \simeq (\det \mathbf{K})^{-(p-1)/2} \\ &= \int \prod_{(\alpha\beta)}^n dq_{\alpha\beta} \delta \left(q_{\alpha\beta} - \sum_i^N w_i S_i^{\alpha} S_i^{\beta} \right) \exp \left(-\frac{p-1}{2} \text{Tr}_n \ln \{ \mathbf{I} - K\mathbf{T}_1 (\mathbf{I} + \mathbf{Q}) \} \right), \end{aligned} \quad (\text{B8})$$

where $(\alpha\beta)$ denotes a summation over α and $\beta (\neq \alpha)$. Here, we used the fact that the diagonal element $K_{\alpha\alpha}$ of the matrix \mathbf{K} equals to $1 - K\mathbf{T}_1$ and the off-diagonal

element becomes

$$K_{\alpha\beta} = -K\mathbf{T}_1 \sum_i^N w_i S_i^{\alpha} S_i^{\beta} \equiv -K\mathbf{T}_1 q_{\alpha\beta}, \quad (\text{B9})$$

which was expressed in terms of the spin glass order parameter $q_{\alpha\beta}$. The matrix \mathbf{Q} consists of zeros along the diagonal elements and the off-diagonal elements $q_{\alpha\beta}$. Tr_n

is the trace of the $n \times n$ matrix. Using the Fourier representation of the δ function in Eq. (B8) with integral variable $\tilde{r}_{\alpha\beta}$, we can obtain the result:

$$\begin{aligned} \langle \langle Z^n \rangle_\xi \rangle_K = & \text{Tr} \int \prod_{\alpha}^n dm_{\alpha}^1 \prod_{(\alpha\beta)}^n dq_{\alpha\beta} \prod_{(\alpha\beta)}^n d\tilde{r}_{\alpha\beta} \exp \left[NK\mathbf{T}_1 \sum_{\alpha}^n \left\{ -\frac{1}{2}(m_{\alpha}^1)^2 + \sum_i^N m_{\alpha}^1 (w_i \xi_i^1 S_i^{\alpha}) \right\} \right] \\ & \times \exp \left[iNK\mathbf{T}_1 \sum_{(\alpha\beta)}^n \tilde{r}_{\alpha\beta} \left(q_{\alpha\beta} - \sum_i^N w_i S_i^{\alpha} S_i^{\beta} \right) \right] \exp \left(-\frac{p-1}{2} \text{Tr}_n \ln \{ (1 - K\mathbf{T}_1)\mathbf{I} - K\mathbf{T}_1\mathbf{Q} \} \right). \end{aligned} \quad (\text{B10})$$

The many-spin coupled problem is transferred to an average over a decoupled problem of single spins. For

this simplification, we deal with inter-replica couplings in the single-spin problem. So we have

$$\langle \langle Z^n \rangle_\xi \rangle_K = \text{Tr} \int \prod_{\alpha}^n dm_{\alpha}^1 \prod_{(\alpha\beta)}^n dq_{\alpha\beta} \prod_{(\alpha\beta)}^n d\tilde{r}_{\alpha\beta} \exp \{ -NG(m_{\alpha}^1, q_{\alpha\beta}, \tilde{r}_{\alpha\beta}) \}, \quad (\text{B11})$$

where

$$G(m_{\alpha}^1, q_{\alpha\beta}, \tilde{r}_{\alpha\beta}) \equiv \frac{1}{2} K\mathbf{T}_1 \sum_{\alpha}^n (m_{\alpha}^1)^2 - iK\mathbf{T}_1 \sum_{(\alpha\beta)}^n \tilde{r}_{\alpha\beta} q_{\alpha\beta} + \frac{a}{2} \text{Tr}_n \ln \{ (1 - K\mathbf{T}_1)\mathbf{I} - K\mathbf{T}_1\mathbf{Q} \} - \ln \text{Tr}_i \exp(\tilde{\mathcal{H}}_i) \quad (\text{B12})$$

with

$$\tilde{\mathcal{H}}_i \equiv NK\mathbf{T}_1 \left(\sum_{\alpha}^n m_{\alpha}^1 (w_i \xi_i^1 S_i^{\alpha}) - i \sum_{(\alpha\beta)}^n \tilde{r}_{\alpha\beta} (w_i S_i^{\alpha} S_i^{\beta}) \right). \quad (\text{B13})$$

Here we used $p-1 \simeq p = aN$ and the trace Tr_i is now at a single spin site.

In the thermodynamic limit ($N \rightarrow \infty$) the integrals can be performed by the steepest descent method:

$$\int dy \exp \{ -NG(y) \} \simeq \int dy \exp \left\{ -NG(y_0) - \frac{1}{2} NG''(y_0) (y - y_0)^2 + \dots \right\}, \quad (\text{B14})$$

where $G'(y_0) = 0$ determines a saddle point y_0 . The Gaussian term can be ignored for $N \rightarrow \infty$, provided that $G'''(y_0) > 0$. Otherwise, the resulting integral diverges and the saddle point procedure fails. Assuming $G'''(y) > 0$, we replace y with their stationary value.

Therefore, we can define three order parameters as follows: i) m_{α}^{μ} represents the extent to which the μ -th pattern of memory ξ^{μ} and the α -th state of the system S^{α}

overlap with each other;

$$m_\alpha^\mu = \sum_i^N w_i \langle \xi_i^\mu S_i^\alpha \rangle. \quad (\text{B15})$$

ii) $q_{\alpha\beta}$ is the spin glass order parameter representing the extent to which the two states α and β of the replica

overlap each other;

$$q_{\alpha\beta} = \sum_i^N w_i \langle S_i^\alpha S_i^\beta \rangle. \quad (\text{B16})$$

iii) $r_{\alpha\beta}$ represents the extent to which the two different m_α^μ 's overlap each other;

$$r_{\alpha\beta} \equiv \frac{N}{p} \sum_{\mu=2}^p m_\alpha^\mu m_\beta^\mu = \frac{1}{a} \sum_{\mu=2}^p q_{\alpha\beta} = -\frac{2i}{aK\mathbf{T}_1} \tilde{r}_{\alpha\beta}. \quad (\text{B17})$$

So, $r_{\alpha\beta}$ can be understood as the sum of the effects of non-retrieved patterns. Here, the average is evaluated through $\langle A \rangle \equiv \text{Tr}_i A \exp \tilde{\mathcal{H}}_i / \text{Tr}_i \exp \tilde{\mathcal{H}}_i$.

Therefore, the free energy becomes

$$n\beta f = \frac{1}{2} K \mathbf{T}_1 \sum_\alpha^n (m_\alpha^1)^2 + \frac{1}{2} a K^2 \mathbf{T}_2 \sum_{(\alpha\beta)}^n r_{\alpha\beta} q_{\alpha\beta} + \frac{a}{2} \text{Tr}_n \ln \{ (1 - K \mathbf{T}_1) \mathbf{I} - K \mathbf{T}_1 \mathbf{Q} \} - \ln \text{Tr}_i \exp(\tilde{\mathcal{H}}_i), \quad (\text{B18})$$

where

$$\tilde{\mathcal{H}}_i = NK \mathbf{T}_1 \sum_\alpha^n m_\alpha^1 (w_i \xi_i^1 S_i^\alpha) + \frac{1}{2} a N K^2 \mathbf{T}_2 \sum_{(\alpha\beta)}^n r_{\alpha\beta} (w_i S_i^\alpha S_i^\beta). \quad (\text{B19})$$

Appendix C: Replica-symmetric (RS) and RS breaking solutions

Here we take the replica-symmetric (RS) assumption to set $q_{\alpha\beta} = q$ and $r_{\alpha\beta} = r$ for all $\alpha \neq \beta$, and $m_\alpha^1 = m$

for all α . Then the free energy is given by

$$\beta f = \lim_{n \rightarrow 0} \frac{1}{n} \left[\frac{1}{2} K \mathbf{T}_1 n m^2 + \frac{1}{2} K^2 \mathbf{T}_2 n(n-1) a r q + \frac{a}{2} \text{Tr}_n \ln \{ (1 - K \mathbf{T}_1) \mathbf{I} - K \mathbf{T}_1 \mathbf{Q} \} - \ln \text{Tr}_i \exp(\tilde{\mathcal{H}}') \right] \quad (\text{C1})$$

with the effective Hamiltonian $\tilde{\mathcal{H}}'$

$$\tilde{\mathcal{H}}' \equiv NK \mathbf{T}_1 w_i m \xi_i^1 \sum_\alpha^n S_i^\alpha + \frac{1}{2} NK^2 \mathbf{T}_2 w_i a r \sum_{(\alpha\beta)}^n S_i^\alpha S_i^\beta \quad (\text{C2})$$

To calculate the trace of the third term of Eq. (C1), it should be noted that the eigenvectors of the matrix \mathbf{Q}

are, first, the uniform one ${}^t(1, 1, \dots, 1)$ and, second, the form of ${}^t(1, 0, \dots, 0, -1, 0, \dots, 0)$. So, the eigenvalue of the first eigenvector of the matrix $(1 - K \mathbf{T}_1) \mathbf{I} - K \mathbf{T}_1 \mathbf{Q}$ is $1 - K \mathbf{T}_1 - (n-1) K \mathbf{T}_1 q$ (no degeneracy), and the eigenvalue of the second eigenvector is $1 - K \mathbf{T}_1 + K \mathbf{T}_1 q$ (degeneracy $n-1$). Thus, in the limit $n \rightarrow 0$,

$$\begin{aligned} \frac{1}{n} \text{Tr}_n \ln \{ (1 - K \mathbf{T}_1) \mathbf{I} - K \mathbf{T}_1 \mathbf{Q} \} &= \frac{1}{n} \ln(1 - K \mathbf{T}_1 - (n-1) K \mathbf{T}_1 q) + \frac{n-1}{n} \ln(1 - K \mathbf{T}_1 + K \mathbf{T}_1 q) \\ &\xrightarrow{n \rightarrow 0} \ln(1 - K \mathbf{T}_1 + K \mathbf{T}_1 q) - \frac{K \mathbf{T}_1 q}{1 - K \mathbf{T}_1 + K \mathbf{T}_1 q}. \end{aligned} \quad (\text{C3})$$

When we put $\xi^1 = 1$, we finally obtain

$$\beta f = \frac{1}{2}K^2\mathbf{T}_2ar(1-q) + \frac{1}{2}K\mathbf{T}_1m^2 + \frac{a}{2}\left[\ln(1-K\mathbf{T}_1+K\mathbf{T}_1q) - \frac{K\mathbf{T}_1q}{1-K\mathbf{T}_1+K\mathbf{T}_1q}\right] - \int \mathcal{D}z \frac{1}{N} \sum_{i=1}^N \ln \left[2 \cosh \eta_i(z)\right]. \quad (\text{C4})$$

where $\int \mathcal{D}z \dots \equiv \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dz \exp\{-\frac{1}{2}z^2\} \dots$, and $\eta_i(z) \equiv K\mathbf{T}_1(z\sqrt{Nw_iar} + Nw_im)$.

We can determine m , r , and q by imposing the condition that f resumes the stable extrema when they are the RS solutions. From this extremal condition, we can obtain the self-consistent equations of m , q , and r as follows:

$$m = \int \mathcal{D}z \sum_{i=1}^N w_i \tanh \eta_i(z) \quad (\text{C5})$$

$$q = \int \mathcal{D}z \sum_{i=1}^N w_i \tanh^2 \eta_i(z) \quad (\text{C6})$$

$$r = \frac{q}{(1-K\mathbf{T}_1+K\mathbf{T}_1q)^2}. \quad (\text{C7})$$

The Almeida-Thouless (AT) line, i.e., the condition satisfying $G''(y_0) = 0$ in Eq. (B14), is simply given by [4, 19]

$$(1-K\mathbf{T}_1+K\mathbf{T}_1q)^2 - K^2\mathbf{T}_2a \int \mathcal{D}z \sum_{i=1}^N Nw_i^2 \text{sech}^4 \eta_i(z) = 0. \quad (\text{C8})$$

As shown in Fig. 1, the AT lines for various γ values locate in very low temperature region near zero so that even the SG and M phases become stable under replica symmetry.

We consider a particular case $T \rightarrow 0$ (i.e., $\mathbf{T}_l \rightarrow 1$ ($l = 1, 2$)). In this limit the “tanh” reduces to a step function

$$\tanh \eta_i(z) \rightarrow \text{sgn}(\eta_i(z)), \quad (\text{C9})$$

where $\eta_i(z) = K(z\sqrt{Nw_iar} + Nw_im)$. Then the equation for m becomes

$$\begin{aligned} m &= \int \mathcal{D}z \sum_{i=1}^N w_i \text{sgn}(\eta_i(z)) \\ &= \sum_{i=1}^N w_i \frac{2}{\sqrt{2\pi}} \int_0^{\sqrt{\frac{Nw_i}{ar}}m} dz e^{-\frac{1}{2}z^2} \\ &= \sum_{i=1}^N w_i \text{erf}\left(\sqrt{\frac{Nw_i}{2ar}}m\right), \end{aligned} \quad (\text{C10})$$

where $\text{erf}(\dots)$ means the error function. The parameter q approaches one, i.e., $q \rightarrow 1$ in the zero temperature limit. A simple equation in this limit is readily obtained

from $\partial(\beta f)/\partial r = 0$:

$$\begin{aligned} NK(1-q) &= \sum_{i=1}^N \sqrt{\frac{Nw_i}{ar}} \int \mathcal{D}z z \text{sgn}(\eta_i(z)) \\ &= \sum_{i=1}^N \sqrt{\frac{Nw_i}{ar}} \frac{2}{\sqrt{2\pi}} \int_{\sqrt{\frac{Nw_i}{ar}}m}^{\infty} dz z e^{-\frac{1}{2}z^2} \\ &= \sum_{i=1}^N \sqrt{\frac{2Nw_i}{\pi ar}} \exp\left(-\frac{Nw_i}{2ar}m^2\right). \end{aligned} \quad (\text{C11})$$

Therefore, at zero temperature, m , q and r are obtained as

$$m = \sum_{i=1}^N w_i \text{erf}\left(\sqrt{\frac{Nw_i}{2ar}}m\right) \quad (\text{C12})$$

$$q = 1 - \frac{1}{NK} \sum_{i=1}^N \sqrt{\frac{2Nw_i}{\pi ar}} \exp\left(-\frac{Nw_i}{2ar}m^2\right) \quad (\text{C13})$$

$$r = \frac{q}{(1-K+Kq)^2}. \quad (\text{C14})$$

Appendix D: Dependence of the phase boundaries on degree exponent

As γ is decreased, the $T - a$ phase diagram in Fig. 1 undergoes drastic changes. Especially, the R phase introduces into the region of the SG phase, but it also raises the boundary of the P phase to a high-temperature region as γ is decreased. As γ approach 2.0, it is shown that the region of R phase becomes broader whereas the SG phase shrinks and eventually disappears. Such changes of the phase diagram in the $T - a$ space can be obtained by the leading term in the expansion of the right hand side of Eq. (C6) with $m = 0$, i.e., $q \simeq arK^2\mathbf{T}_2X$ with $X \equiv N \sum_{i=1}^N w_i^2$, under the condition $N^{1/(\gamma-1)} \ll 1$. From this, we obtain the glass transition temperature to be $T_g \simeq 1/K \tanh^{-1}(1/K(1 + \sqrt{aX}))$. For the ER network ($\gamma \rightarrow \infty$) with $K = 5$ and $a = 0$, we obtain $T_g \simeq 1.013$ as shown in Fig. 1(a) of the main paper. However, as $\gamma \rightarrow 2$, the condition $N^{1/(\gamma-1)} \ll 1$ cannot be fulfilled as $N \rightarrow \infty$, thus there exists no phase boundary between the P and SG phases and the R phase intrudes into the region between the two phases.

Appendix E: Dependence of the error rate on degree exponent

Fig. 2 shows the error rate $n_e \equiv (1-m)/2$ as a function storage capacity a , obtained by Eqs. (C12-C14) for various γ values at zero temperature. As shown in Eq.(C12), m has a γ dependence in terms of w_i . For the case $\gamma \rightarrow \infty$, Nw_i becomes unity and Eq.(C12) reduces to $m = \text{erf}(m/\sqrt{2ar})$, by which m suddenly becomes zero when a is larger than a specific value called the critical capacity $a_c = a_c(\gamma)$. For the case $\gamma \rightarrow 2+\epsilon$ ($\epsilon \ll 1$), however, $Nw_i \sim N^{1/(1+\epsilon)}i^{-1/(1+\epsilon)} \gg 1$ as $N \rightarrow \infty$, by which m becomes nonzero even for sufficiently large value of a and it is thus impossible to define the critical capacity a_c . Such a γ -dependence of m determines the behavior of the error rate n_e , as shown in Fig. 2.

Appendix F: Comparison between the Chung-Lu model and the static model

We show the phase diagram in the space $(T-a)$ and the error rate $n_e = (1-m)/2$ as a function of storage capacity a for the static model. First, Figs. 3(a) and (b) show the phase diagrams for the static model with (a) $\gamma = 2.35$ and (b) $\gamma = 2.01$, which correspond to the Figs. 3(e) and (f) of the main paper for the CL model. These phase diagrams for the static model are similar qualitatively to those for the CL model, but the transitions temperature between P and R phases obtained for the static model is somewhat larger than the one for the CL model.

Fig. 4 is the plot of the error rate $n_e \equiv (1-m)/2$ vs storage capacity a at zero temperature for the two different γ values for the static model. This figure corresponds

to Fig. 2(b) for the CL model. Whereas the SG phase of the CL model remains on the axis $T = 0$ when $\gamma \simeq 2.04$, as shown in Fig. 2(b), the SG phase of the static model given in Fig. 4 remains on the axis $T = 0$ when $\gamma \simeq 2.35$.

Figs. 5 show the comparison of the error rate for the two different models, the CL and the static models as a function of storage rate a for different degree exponent values (a) $\gamma = 2.01$ and (b) 2.04 at $T = 0$. These figures enable us to see qualitatively how the error rate depends on the degree-degree correlation. We recall that the disassortative degree-degree correlation is present for the static model in the region $2 < \gamma < 3$, but absent for the CL model. One can see that the degree-degree correlation reduces somewhat the error rate. Accordingly one may infer that the error rate of realistic brain networks, in which the degree-degree correlation is present, should be smaller than those obtained from the CL model analytically.

Appendix G: Hopfield model simulations on real neural networks

We check the error rate n_e of the Hopfield model on several real neural networks. For this purpose, we chose three real neural networks: the networks for Macaque monkey 1 (fve30; $N = 30$, $L = 311$) [37], Macaque monkey 2 (macaque47; $N = 47$, $L = 505$) [38] and cat (CI-Jctx; $N = 52$, $L = 818$) [39]. Those networks are constructed based on the connectivity network data sets [40]. For the cat network, all links with nonzero elements (1,2 and 3) of the adjacency matrix were regarded as connected.

Our simulations are performed following the random updating method of the Hopfield model on those real neural networks: An excitation pattern of a neural network is denoted by $\{\xi_i^\mu\}$, where i is the index of the neuron, $\mu (= 1, \dots, p)$ is the index of the excitation pattern, and p is the number of existing patterns, equivalent to the number of stored memories. The value of ξ_i^μ is either +1 or -1. An initial spin configuration of each neural network is set to be the first pattern, i.e., $S_i(t=1) = \xi_i^{\mu=1}$ for each node i . Once this spin configuration is set up, at the next step $t = 2$, a neuron i is randomly chosen, its Ising spin state is updated asynchronously as

$$S_i(t+1) = \text{sgn}\left(\sum_{j \in \text{n.n. of } i} J_{ij} S_j(t)\right), \quad (\text{G1})$$

where the synapse efficacy J_{ij} is taken as $J_{ij} = (1/K) \sum_{\mu=1}^p \xi_i^\mu \xi_j^\mu$ following the Hebb's rule, and $K (\equiv 2L/N)$ is the mean degree of a given real neural network. If $\sum_j J_{ij} S_j(t)$ becomes zero, then $S_i(t+1)$ takes +1 definitely. This updating is repeated until the system reaches a stable fixed point, which is supposed to be the energy minimum state of the system. Thus the retrieved pattern is the spin configuration in the energy minimum state. The error rate $n_e (\equiv (1-m)/2)$ is finally calculated using the formula $m = (1/N) \sum_i \xi_i^1 S_i(t \rightarrow \infty)$

with $\mu = 1$. This whole procedure completes one run. We perform 10^4 different runs with 10^4 different ensemble of $\{\xi_i^\mu\}$, and then obtain the error rate n_e over those ensemble.

Fig. 6 shows the degree distributions (left column) and the error rates (right column) of those three different real neural networks. The degree distributions of three networks are heavy-tailed, but the degree exponents cannot

be determined precisely due to the small system sizes of those networks. To compare the error rates obtained from the simulations with the ones from the theory, we draw the error rates obtained from both methods together in the right column of Fig. 6. Indeed, with appropriate weight $w_i(\gamma)$ near $\gamma = 2.0$, the two error rates behave similarly to each other as a function of the storage capacity.

-
- [1] S. A. Josselyn, S. Köhler, and P. W. Frankland, *Finding The Engram*, Nat. Rev. Neurosci. **16**, 521-534 (2015).
 - [2] J. J. Hopfield, *Neural networks and physical systems with emergent collective computational abilities*, Proc. Natl. Acad. Sci. USA **79**, 2554-2558 (1982).
 - [3] D. J. Amit, H. Gutfreund, and H. Sompolinsky, *Storing infinite numbers of patterns in a spin-glass model of neural networks*, Phys. Rev. Lett. **55**, 1530-1533 (1985).
 - [4] D. J. Amit, H. Gutfreund, and H. Sompolinsky, *Statistical mechanics of neural networks near saturation*, Ann. Phys. (NY) **173**, 30-67 (1987).
 - [5] D. J. Amit, *Modeling Brain Function: The World of Attractor Neural Networks* (Cambridge University Press, 1989).
 - [6] E. Bullmore, and O. Sporns, *Complex brain networks: graph theoretical analysis of structural and functional systems*, Nat. Rev. Neurosci. **10**, 186-198 (2009).
 - [7] M. P. van den Heuvel, and O. Sporns, *Network hubs in the human brain*, Trends. Cogn. Sci. **17**, 683-696 (2013).
 - [8] A. -L. Barabási, and R. Albert, *Emergence of scaling in random networks*, Science **286**, 509-512 (1999).
 - [9] V. M. Eguíluz, D. R. Chialvo, G. A. Cecchi, M. Baliki, and A. V. Apkarian, *Scale-free Brain Functional Networks*, Phys. Rev. Lett. **94**, 018102 (2005).
 - [10] M. P. van den Heuvel, C. J. Stam, M. Boersma, and H. E. H. Pol, *Small-world and scale-free organization of voxel-based resting-state functional connectivity in the human brain*, NeuroImage **43**, 528-539 (2008).
 - [11] L. K. Gallos, H. A. Makse, and M. Sigman, *A small world of weak ties provides optimal global integration of self-similar modules in functional brain networks*, Proc. Natl. Acad. Sci. USA **109**, 2825-2830 (2012).
 - [12] M. Rubinov, and E. Bullmore, *Schizophrenia and abnormal brain network hubs*. Dialogues Clin. Neurosci. **15**, 339-349 (2013).
 - [13] P. Steullet, J. H. Cabungcal, A. Monin, D. Dwir, P. O'Donnell, M. Cuenod, and K. Q. Do, *Redox dysregulation, neuroinflammation, and NMDA receptor hypofunction: A central hub in schizophrenia pathophysiology?* Schizophr. Res. in press (2014), <http://dx.doi.org/10.1016/j.schres.2014.06.021>
 - [14] S. Achard, C. Delon-Martin, P. E. Vértes, F. Renard, M. Schenck, F. Schneider, C. Heinrich, S. Kremer, and E. T. Bullmore, *Hubs of brain functional networks are radically reorganized in comatose patients*, Proc. Natl. Acad. Sci. USA **109**, 20608-20613 (2012).
 - [15] D. Stauffer, A. Aharony, L. da Fontoura Costa, and J. Adler, *Efficient Hopfield pattern recognition on a scale-free neural network*, Eur. Phys. J. B **32**, 395-399 (2003).
 - [16] J. J. Torres, M. A. Muñoz, J. Marro, and P. L. Garrido, *Influence of topology on the performance of a neural network*, Neurocomputing **58-60**, 229-234 (2004).
 - [17] F. Chung, and L. Lu, *Connected components in random graphs with given expected degree sequences*, Ann. Comb. **6**, 125-145 (2002).
 - [18] Y. S. Cho, J. S. Kim, J. Park, B. Kahng, and D. Kim, *Percolation transitions in scale-free networks under the Achlioptas process*, Phys. Rev. Lett. **103**, 135702 (2009).
 - [19] J. R. L. de Almeida, and D. J. Thouless, *Stability of the Sherrington-Kirkpatrick solution of a spin glass model*, J. Phys. A: Math. Gen. **11**, 983-990 (1978).
 - [20] T. M. Bartol Jr., C. Bromer, J. Kinney, M. A. Chirillo, J. N. Bourne, K. M. Harris, and T. J. Sejnowski, *Nanoconnectomic upper bound on the variability of synaptic plasticity*, eLife **4**, e10778 (2015).
 - [21] M. Mézard, G. Parisi, and M. A. Virasoro, *Spin Glass Theory and Beyond* (World Scientific, 1987).
 - [22] H. Nishimori, *Statistical Physics of Spin Glasses and Information Processing: An Introduction* (Oxford University Press, 2001).
 - [23] K. -I. Goh, B. Kahng, and D. Kim, *Universal behavior of load distribution in scale-free networks*, Phys. Rev. Lett. **87**, 278701 (2001).
 - [24] J. -S. Lee, K. -I. Goh, B. Kahng, and D. Kim, *Intrinsic degree-correlations in the static model of scale-free networks*, Eur. Phys. J. B. **49**, 231-238 (2006).
 - [25] D. -H. Kim, G. J. Rodgers, B. Kahng, and D. Kim, *Spin-glass phase transition on scale-free networks*, Phys. Rev. E **71**, 056115 (2005).
 - [26] D. -H. Kim, *Inverse transitions in a spin-glass model on a scale-free network*, Phys. Rev. E **89**, 022803 (2014).
 - [27] I. P. Castillo, B. Wemmenhove, J. P. L. Hatchett, A. C. C. Coolen, N. S. Skantzos, and T. Nikolettopoulos, *Analytic solution of attractor neural networks on scale-free graphs*, J. Phys. A: Math. Gen. **37**, 8789-8799 (2004).
 - [28] P. Erdős, and A. Rényi, *On random graphs I*, Publ. Math.-Debr. **6**, 290-297 (1959).
 - [29] P. Erdős, and A. Rényi, *On the evolution of random graphs*, Publ. Math. Inst. Hung. Acad. Sci. **5**, 17-61 (1960).
 - [30] B. Bollobás, *Random Graphs* (Cambridge University Press, 2001).
 - [31] L. Viana, and A. J. Bray, *Phase diagrams for dilute spin glasses*, J. Phys. C: Solid State Phys. **18**, 3037-3051 (1985).
 - [32] M. Mézard, and G. Parisi, *Mean-field theory of randomly frustrated systems with finite connectivity*, Europhys. Lett. **3**, 1067-1074 (1987).
 - [33] I. Kanter, and H. Sompolinsky, *Mean-field theory of spin-glasses with finite coordination number*, Phys. Rev. Lett. **58**, 164 (1987).
 - [34] P. Mottishaw, and C. De Dominicis, *On the stability of*

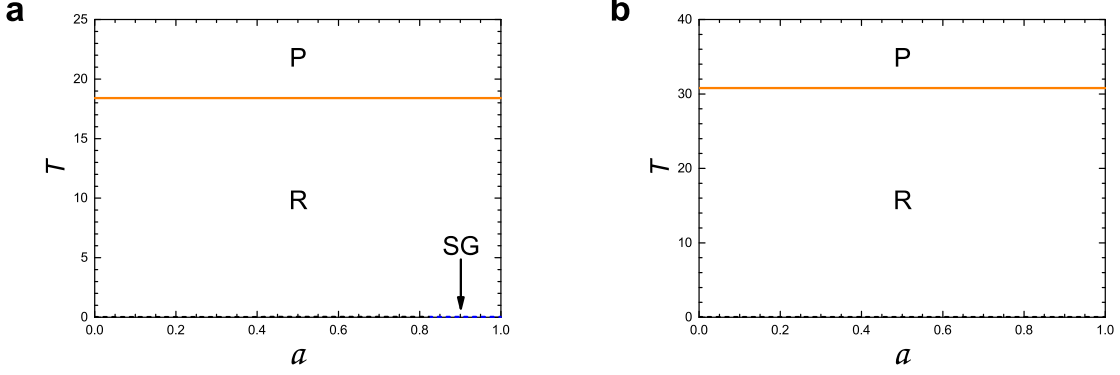


FIG. 3. Phase diagrams of the Hopfield model in the plane of (T, a) on the static model of SF networks with (a) $\gamma = 2.35$ and (b) $\gamma = 2.01$. Those phase diagrams correspond to Fig. 1(e) and (f) of the main paper for the CL model, respectively. **P** represents paramagnetic phase, **SG** spin glass phase, and **R** retrieval phase. The SG phase remains only on the axis $T = 0$ when $\gamma = \gamma_c \simeq 2.35$, then the R phase spans the entire region of a at the lower temperatures. Thus, the system becomes more memory retrieval. To obtain the phase boundary, numerical calculations were performed for the static model with plugging $N = 1000$ and $K = 5.0$ into the formulas. Solid and dotted curves indicate the second-order and the first-order transitions, respectively. Note that black dotted line in each panel near the $T = 0$ line represents the Almeida-Thouless line [19]. Thus, replica-symmetric solution is valid over almost the entire region of the phase space.

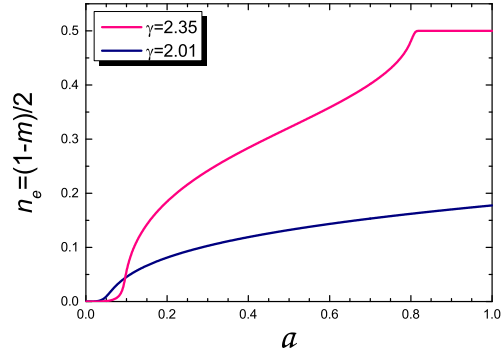


FIG. 4. Plot of the error rate $n_e \equiv (1 - m)/2$ vs storage capacity a for $\gamma = 2.01$ and 2.35 for the static model at $T = 0$, which corresponds to Fig. 2(b) of the main paper for the CL model. Here, numerical values are obtained using $N = 1000$ and $K = 5.0$.

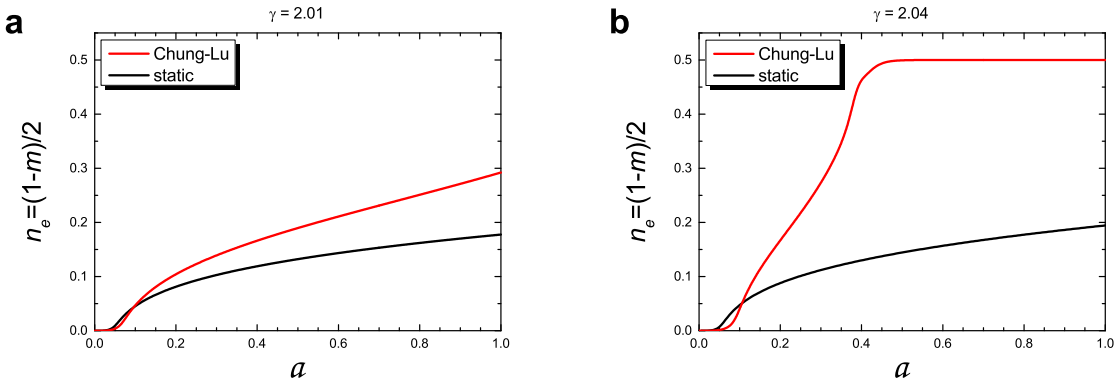


FIG. 5. Plot of the error rate n_e vs storage capacity a for (a) $\gamma = 2.01$ and (b) $\gamma = 2.04$ for the CL and the static model at $T = 0$. These figures provide the comparison of the error rate between for the CL model and for the static model. Here, numerical values are obtained using $N = 1000$ and $K = 5.0$.

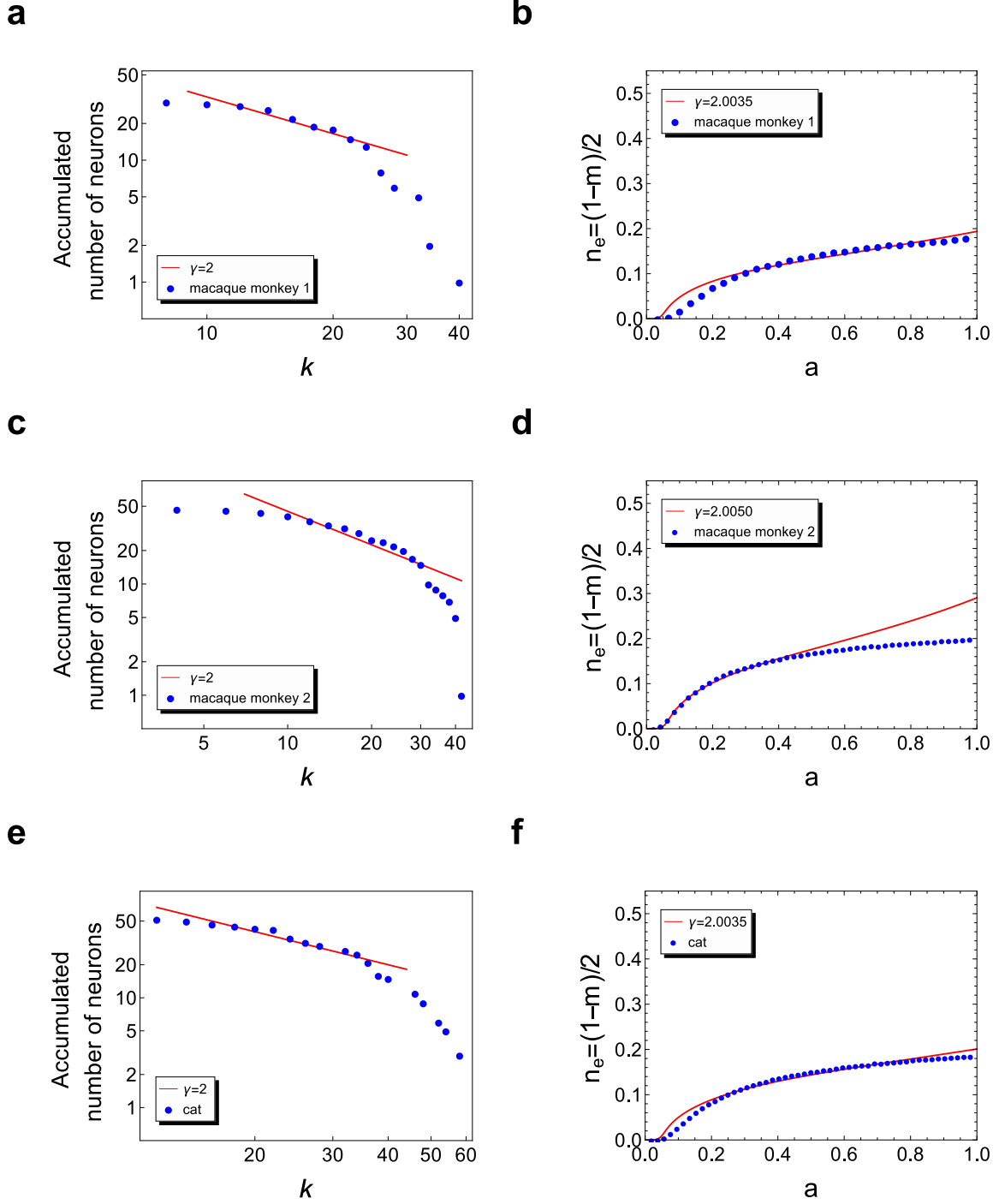


FIG. 6. Plots of the number of neurons (nodes) with degree larger than k vs degree k of the real neural networks (left column) and of their error rates n_e vs storage capacity a (right column): (a) and (b) for the visual cortex network of macaque monkey, with $N = 30$ and $L = 311$ [37]. (c) and (d) for the corticocortical connectivity network in the visual and sensorimotor area of macaque monkey, with $N = 47$ and $L = 505$ [38]. (e) and (f) for the cortex network of cat, with $N = 52$ and $L = 818$ [39]. The red lines of (a), (c) and (e) are guidelines with slope -1 for eye drawn to compare the data points with the scale-freeness of $\gamma = 2.0$. The red curves of (b), (d) and (f) are drawn to compare the simulation data (\bullet) with the analytic solution (red curve) using Eq.(40) under the same conditions of N and L . Here, the γ values we used for w_i in Eq. (40) were 2.0035 (b), 2.0050 (d), and 2.0035 (f), respectively.

- randomly frustrated systems with finite connectivity*, J. Phys. A: Math. Gen. **20**, L375-L379 (1987).
- [35] K. Y. Wong, and D. Sherrington, *Intensively connected spin glasses: towards a replica-symmetry-breaking solution of the ground state*, J. Phys. A: Math. Gen. **21**, L459-L466 (1988).
 - [36] R. Monasson, *Optimization problems and replica symmetry breaking in finite connectivity spin glasses*, J. Phys. A: Math. Gen. **31**, 513-529 (1998).
 - [37] D. J. Felleman, and D. C. Van Essen, *Distributed hierarchical processing in the primate cerebral cortex*, Cereb. Cortex **1**, 1-47 (1991).
 - [38] C. J. Honey, R. Kötter, M. Breakspear, and O. Sporns, *Network structure of cerebral cortex shapes functional connectivity on multiple time scales*, Proc. Natl. Acad. Sci. USA **104**, 10240-10245 (2007).
 - [39] J. W. Scannell, G. A. P. C. Burns, C. C. Hilgetag, M. A. O'Neil, and M. P. Young, *The connectional organization of the cortico-thalamic system of the cat*, Cereb. Cortex **9**, 277-299 (1999).
 - [40] The functional brain connectivity data sets of macaque monkey and cat can be downloaded from <https://sites.google.com/site/bctnet/datasets/>